

GeneNet in 2005

**E. A. Ananko*, N. L. Podkolodny, I. L. Stepanenko, O. A. Podkolodnaya,
D. A. Rasskazov, D. S. Miginsky, V. A. Likhoshvai, A. V. Ratushny,
N. N. Podkolodnaya and N. A. Kolchanov**

Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences,
Lavrentiev Avenue 10, Novosibirsk 630090, Russia

Received September 15, 2004; Revised and Accepted October 8, 2004

ABSTRACT

The GeneNet system is designed for collection and analysis of the data on gene and metabolic networks, signal transduction pathways and kinetic characteristics of elementary processes. In the past 2 years, the GeneNet structure was considerably improved: (i) the current version of the database is now implemented using ORACLE9i; (ii) the capacities to describe the structure of the protein complexes and the interactions between the units are increased; (iii) two tables with kinetic constants and more detailed descriptions of certain reactions were added; and (iv) a module for kinetic modeling was supplemented. The current SRS release of the GeneNet database contains 37 graphical maps of gene networks, as well as descriptions of 1766 proteins, 1006 genes, 241 small molecules and 3254 relationships between gene network units, and 552 kinetic constants. Information distributed between 16 inter-linked tables was obtained by annotating 1980 journal publications. SRS release of the GeneNet database, the graphical viewer and the modeling section are available at <http://www.mgs.bionet.nsc.ru/mgs/gnw/genenet/>.

INTRODUCTION

Systematic arrangement and analysis of a variety of data subsets present a challenge. The GeneNet system (1) was developed to respond to the challenge. The widely available specialized databases, such as KEGG (2), BioCarta (<http://www.biocarta.com/index.asp>), BIND (3), TRANSPATH (4) and MetaCyc (5), among others, are concerned with metabolic or signal transduction pathways. The GeneNet workbench is more versatile, enabling the description of functions and regulation of complicated biological systems on the basis

of the manifold data on the expression regulation of genes and changes in the concentration of their products, the macromolecular interactions, enzymatic reactions, the effects of external agents, signal transduction pathways, to name a few. The GeneNet diagrams represent mainly the structural–functional organization of the gene network (1) that controls particular processes in eukaryotes (6). The GeneNet's task is not to give a detailed description of protein–protein interactions or of the gene or protein structure; all this is available from the other established databases. The intention is to describe, in more detail, the relationships between the gene network units and regulatory influences on the relationships (7). The information from the GeneNet database is further used in the developing of kinetic computer models of various biological processes (8).

GeneNet MODULES

The GeneNet system consists of the following functional modules:

- (i) a database that compiles information on the structural and functional organization of the gene and metabolic networks, their elementary units (proteins, genes, RNAs, small molecules, etc.) and elementary interactions between the units (1);
- (ii) a graphical viewer that allows display of the structure of the gene network and interactions of its units as a two-dimensional graph;
- (iii) a section of computer modeling of gene networks;
- (iv) software for analysis of the graph structure of gene network and its functional characteristics (9);
- (v) a graphical editor enabling the construction of new diagrams and to modify the existing ones, to add new information to the database.

It should be noted that the user can request the software for analysis and the graphical editor in the GeneNet licensed version only. For this reason, the descriptions of these tools are omitted here. Furthermore, the licensed GeneNet version

*To whom correspondence should be addressed. Tel: +7 3832 333119; Fax: +7 3832 331278; Email: eananko@bionet.nsc.ru

The online version of this article has been published under an open access model. Users are entitled to use, reproduce, disseminate, or display the open access version of this article for non-commercial purposes provided that: the original authorship is properly and fully attributed; the Journal and Oxford University Press are attributed as the original place of publication with the correct citation details given; if an article is subsequently reproduced or disseminated not in its entirety but only in part or as a derivative work this must be clearly indicated. For commercial re-use permissions, please contact journals.permissions@oupjournals.org.

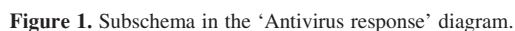
Currently, information on the elementary structures and interactions are stored as a relational database with nested tables (their total number is 89); the gene and metabolic network diagrams and the database structure are described in the XML format.

The internal format of the GeneNet database was considerably expanded. In the current version, we can more completely describe interactions between units, e.g. introduce information on the constants and cofactors of enzymatic reactions, subunit structure, protein modification sites, half-life time of RNAs, proteins and small molecules, among others.

The current SRS release of the GeneNet database contains descriptions of 1766 proteins, 1006 genes, 241 small molecules and 3254 relationships between gene network units. 552

The licensed version of the GeneNet database is in the XML format, its number of entries is about two times greater than in the free version.

In 2004, a new GeneNet viewer was developed. In contrast to the preceding version, it can work both on the standard Windows platform and UNIX or LINUX, and it does not require triggering of special programs by a server. Every diagram generated by the GeneNet viewer is a graphical file in the GIF format displaying a two-dimensional graph of a gene network. The information for compiling the diagrams was obtained from the studies on eukaryotes, such as human, mouse, rat, chicken and others. Every diagram is an integration of the data on all the organisms (7). The images of the elementary units (genes, proteins, small molecules) are interactive. By clicking them, one can obtain a description of the unit in a separate window of the Internet browser. Each arrow, which represents interaction between the units (7), has an interactive point. Information on interaction from the GeneNet database can be obtained by clicking the interactive point. It



should be noted that information on a unit displayed by the GeneNet viewer can be different from that stored in the SRS release. This is because the viewer addresses the relational version of the database containing more information than the SRS release.

Altogether, 37 graphical representations of gene networks can be visualized in the GeneNet free version. The gene networks with the indicated number of elementary units and their interactions are listed in Supplementary File 2.

GeneNet modeling

The GeneNet modeling section allows modification of the parameters of the available gene network models (in the free version, only the 'Cholesterol biosynthesis' and the 'Erythrocyte maturation' models are accessible) and keeping track of the course of changes in unit behavior in various conditions. In addition, 38 mathematical models of gene networks in the SBML format (see Supplementary File 3) are accessible in the GeneNet modeling section. The special software for automated generation of computer models derived from the GeneNet data was developed at the Institute of Cytology and Genetics (Novosibirsk, Russia).

GeneNet: A VERSATILE WORKBENCH

GeneNet can be used for addressing a wide range of problems. Thus, all the reactions and regulatory events involving the gene or protein of interest can be found by a query to the free GeneNet SRS-version. The licensed version allows generation of more complicated queries, e.g. to search for the entire path(s) linking two chosen units, to find units connected not only directly but also by second-, third- and higher-orders with the chosen unit, to identify all the chosen regulatory circuits in the gene network graph and to make other queries.

Furthermore, a system of subschemes is available in GeneNet. It allows the user to display certain processes, e.g. signal transduction pathways, as separate diagrams. This approach enables the simplification of the overall gene network representation and the obtaining of a more detailed representation of a particular process, if necessary. Figure 1 shows the implementation of the subschema for the 'Antiviral response' diagram.

FUTURE EXTENSIONS OF GeneNet

In the next 2 years, we intend to increase considerably the capacities for describing the gene network dynamics, to develop a database for kinetic and other new types of experimental data and also to establish a database of gene network computer models. We also intend to improve the software for

analysis of the gene network graph and other information stored in GeneNet. A web-interface will be provided for the user to make queries to the current relational version of the GeneNet database.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

The authors are grateful to I. V. Lokhova for providing the bibliographical support and to A. N. Fadeeva for translating the manuscript from Russian into English. The work was supported in part by the RFBR (03-07-90359, 03-04-48506), RAS 10.4, the Russian Ministry of Industry, Sciences and Technologies (No. 43.073.1.1.1501), SB RAS (Integration Project No. 119), NIH, USA (No. 2 R01-HG-01539-04A2), the Russian Federal Research Development Program (Contract No. 28/2003).

REFERENCES

1. Ananko,E.A., Podkolodny,N.L., Stepanenko,I.L., Ignatieva,E.V., Podkolodnaya,O.A. and Kolchanov,N.A. (2002) GeneNet: a database on structure and functional organisation of gene networks. *Nucleic Acids Res.*, **30**, 398–401.
2. Kanehisa,M., Goto,S., Kawashima,S., Okuno,Y. and Hattori,M. (2004) The KEGG resource for deciphering the genome. *Nucleic Acids Res.*, **32**, D277–D280.
3. Bader,G.D., Betel,D. and Hogue,C.W. (2003) BIND: the Biomolecular Interaction Network Database. *Nucleic Acids Res.*, **31**, 248–250.
4. Krull,M., Voss,N., Choi,C., Pistor,S., Potapov,A. and Wingender,E. (2003) TRANSPATH: an integrated database on signal transduction and a tool for array analysis. *Nucleic Acids Res.*, **31**, 97–100.
5. Krieger,C.J., Zhang,P., Mueller,L.A., Wang,A., Paley,S., Arnaud,M., Pick,J., Rhee,S.Y. and Karp,P.D. (2004) MetaCyc: a multiorganism database of metabolic pathways and enzymes. *Nucleic Acids Res.*, **32**, D438–D442.
6. Kolchanov,N.A., Ananko,E.A., Kolpakov,F.A., Podkolodnaya,O.A., Ignatieva,E.V., Goryachkovskaya,T.N. and Stepanenko,I.L. (2000) Gene networks. *Mol. Biol. (Moscow)*, **34**, 449–460.
7. Kolpakov,F.A., Ananko,E.A., Kolesov,G.B. and Kolchanov,N.A. (1998) GeneNet: a gene network database and its automated visualization. *Bioinformatics*, **14**, 529–537.
8. Kolchanov,N.A., Nedosekina,E.A., Ananko,E.A., Likhoshvai,V.A., Podkolodny,N.L., Ratushny,A.V., Stepanenko,I.L., Podkolodnaya,O.A., Ignatieva,E.V. and Matushkin,Y.G. (2002) GeneNet database: description and modeling of gene networks. *In Silico Biol.*, **2**, 97–110.
9. Ananko,E.A., Loktev,K.A. and Podkolodny,N.L. (2004) Development and analysis of models of genetic and metabolic networks and signal transduction pathways in the GeneNet system. In Kolchanov,N. and Hofstaedt,R. (eds), *Bioinformatics of Genome Regulation and Structure*. Kluwer Academic Publishers, Boston/Dordrecht/London, pp. 265–272.